UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office

Address : COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

SERIAL NUMBER	FILING DATE	FIRST NAMED APPLICANT	ATTORNEY DOCKET NO.
07/762,762	09/16/91	THOMPSON G	CGNE-69-1

CALGENE, INC.
1920 FIFTH ST.
DAVIS, CA 95616

EXAMINER	
RHODES, P	
ART UNIT	PAPER NUMBER
1804	6

DATE MAILED: 03/27/92

Please find below a communication from the EXAMINER in charge of this application.

Commissioner of Patents.

This application contains sequence disclosures that are encompassed by the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR 1.821(a) (1) and (a) (2). However, this application fails to comply with one or more of the requirements of 37 CFR 1.821 through 1.825 for one or more of the reasons set forth on the attached form "Notice To Comply With Requirements For Patent Applications Containing Nucleotide Sequence And/Or Amino Acid Sequence Disclosures." Wherein attention is directed to paragraph (s) 1 and 11.

Elizabeth C Weimar

ELIZABETH C. WEIMAR
SUPERVISORY PATENT EXAMINER
ART UNIT 184

file copy

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

This application contains sequence disclosures that are encompassed by the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR § 1.821(a)(1) and (a)(2). However, this application fails to comply with one or more of the requirements of 37 CFR §§ 1.821 through 1.825 as follows:

☒ 1. This application clearly fails to comply with the collective requirements of §§ 1.821 through 1.825. Applicant's attention is directed to these regulations, a copy of which is attached.

☐ 2. This application does not conform exclusively to the requirements of §§ 1.821 through 1.825. The non-conforming material should be deleted. § 1.821(b).

☐ 3. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing." § 1.821(c).

☐ 4. This application does contain, as a separate part of the disclosure on paper copy, a "Sequence Listing." However, the "Sequence Listing" does not comply with the requirements of §§ 1.821 through 1.825 as follows:

☐ a. The sequence data does not comply with the symbol and format requirements of paragraphs (b) through (p) of § 1.822. Specifically: _____

_____

☐ b. The "Sequence Listing" does not comply with the location and page requirements of paragraph (a) of § 1.823.

☐ c. The "Sequence Listing" does not comply with the information requirements of paragraph (b) of § 1.823. Specifically: _____

_____

☐ d. Other: _____

☐ 5. The description and/or claims of the patent application mention a sequence that is set forth in the "Sequence Listing" but reference is not properly made to the sequence by use of a sequence identifier as required by § 1.821(d).

☐ 6. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by § 1.821(e).

☐ 7. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the computer readable form does not comply with the requirements of § 1.824. Specifically:

_____

☐ 8. A statement that the content of the paper and computer readable copies are the same has not been submitted as required by § 1.821(f).

☐ 9. The amendment to or replacement of the paper and/or computer readable copies of the "Sequence Listing" does not comply with the requirements of § 1.825(a) through (c).

☐ 10. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable. Applicant must provide a substitute copy of the data in computer readable form accompanied by a statement that the substitute data is identical to that originally filed. § 1.825(d). Specifically: _____

☒ 11. Other: New sequence data in Figures 3, 4A, 4B, 4C, 5, and 8 are not found in the parent applications; thus sequence rules apply to this CIP application. (1114 OG 29)

APPLICANT IS GIVEN ONE MONTH FROM THE DATE OF THIS LETTER WITHIN WHICH TO COMPLY WITH THE ABOVE REQUIREMENTS. Failure to comply with the above requirements will result in ABANDONMENT of the application under 37 CFR 1.821(g). Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR § 1.136. Direct the response to, and any questions about, this notice to the undersigned. A copy of this notice MUST be returned with your response.

☐ _____
For: Manager, Application Processing Division
(703) 308-1202 or 308- ____

☒ P. Rhodes
Examining Group 1804
(703) 308- 3724

REQUIREMENTS FOR APPLICATION DISCLOSURES CONTAINING NUCLEOTIDE AND/OR AMINO ACID SEQUENCES

The full text of the final rule package can be found in the *Federal Register*, 55 Fed. Reg. 18230 (May 1, 1990), and in the *Official Gazette*, 1114 O.G. 29 (May 15, 1990). The text, only, of the regulations is reproduced herein.

§ 1.821 Nucleotide and/or amino acid sequence disclosures in patent applications.

(a)"Nucleotide and/or amino acid sequences" as used in §§ 1.821 through 1.825 is interpreted to mean an unbranched sequence of four or more amino acids or an unbranched sequence of ten or more nucleotides. Branched sequences are specifically excluded from this definition. Nucleotides and amino acids are further defined as follows:

(1) "Nucleotides" are intended to embrace only those nucleotides that can be represented using the symbols set forth in § 1.822(b)(1). Modifications, e.g., methylated bases, may be described as set forth in § 1.822(b), but shall not be shown explicitly in the nucleotide sequence.

(2) "Amino acids" are those L-amino acids commonly found in naturally occurring proteins and are listed in § 1.822(b)(2). Those amino acid sequences containing D-amino acids are not intended to be embraced by this definition. Any amino acid sequence that contains post-translationally modified amino acids may be described as the amino acid sequence that is initially translated using the symbols shown in § 1.822(b)(2) with the modified positions, e.g., hydroxylations or glycosylations, being described as set forth in § 1. 822(b), but these modifications shall not be shown explicitly in the amino acid sequence. Any peptide or protein that can be expressed as a sequence using the symbols in § 1.822(b)(2) in conjunction with a description elsewhere in the "Sequence Listing" to describe, for example, modified linkages, cross links and end caps, non-peptidyl bonds, etc., is embraced by this definition.

(b) Patent applications which contain disclosures of nucleotide and/or amino acid sequences, in accordance with the definition in paragraph (a) of this section, shall, with regard to the manner in which the nucleotide and/or amino acid sequences are presented and described, conform exclusively to the requirements of §§ 1.821 through 1.825.

(c) Patent applications which contain disclosures of nucleotide and/or amino acid sequences must contain, as a separate part of the disclosure on paper copy, hereinafter referred to as the "Sequence Listing," a disclosure of the nucleotide and/or amino acid sequences and associated information using the symbols and format in accordance with the requirements of §§ 1.822 and 1.823. Each sequence disclosed must appear separately in the "Sequence Listing." Each sequence set forth in the "Sequence Listing" shall be assigned a separate identifier written as SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, etc.

(d) Where the description or claims of a patent application discuss a sequence listing that is set forth in the "Sequence Listing" in accordance with paragraph (c) of this section, reference must be made to the sequence by use of the assigned identifier, in the text of the description or claims, even if the sequence is also embedded in the text of the description or claims of the patent application.

(e) A copy of the "Sequence Listing" referred to in paragraph (c) of this section must also be submitted in computer readable form in accordance with the requirements of § 1.824. The

computer readable form is a copy of the "Sequence Listing" and will not necessarily be retained as part of the patent application file. If the computer readable form of a new application is to be identical with the computer readable form of another application of the applicant on file in the Office, reference may be made to the other application and computer readable form in lieu of filing a duplicate computer readable form in the new application. The new application shall be accompanied by a letter making such reference to the other application and computer readable form, both of which shall be completely identified.

(f) In addition to the paper copy required by paragraph (c) of this section and the computer readable form required by paragraph (e) of this section, a statement that the content of the paper and computer readable copies are the same must be submitted with the computer readable form. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(g) If any of the requirements of paragraphs (b) through (f) of this section are not satisfied at the time of filing under 35 U.S.C. 111 or at the time of entering the national stage under 35 U.S.C. 371, applicant has one month from the date of a notice which will be sent requiring compliance with the requirements in order to prevent abandonment of the application. Any submission in response to a requirement under this paragraph must be accompanied by a statement that the submission includes no new matter. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(h) If any of the requirements of paragraphs (b) through (f) of this section are not satisfied at the time of filing, in the United States Receiving Office, an international application under the Patent Cooperation Treaty (PCT) applicant has one month from the date of a notice which will be sent requiring compliance with the requirements, or such other time as may be set by the Commissioner, in which to comply. Any submission in response to a requirement under this paragraph must be accompanied by a statement that the submission does not include new matter or go beyond the disclosure in the international application as filed. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(i) Neither the presence nor the absence of information which is not required under §§ 1.821 through 1.825, in an application shall create any presumption that such information is necessary to satisfy one or more of the requirements of 35 U.S.C.112. Further, the grant of a patent on an application that is subject to the requirements of §§ 1.821 through 1.825 shall constitute a conclusive presumption that said patent complies with the requirements of §§ 1.821 through 1.825.

(j) Envelopes containing only application papers, computer readable forms and fees filed under this section should be marked "Box SEQUENCE."

§ 1.822 Symbols and format to be used for nucleotide and/or amino acid sequence data.

(a) The symbols and format to be used for nucleotide and/or amino acid sequence data shall conform to the requirements of paragraphs (b) through (p) of this section.

(b) The code for representing the nucleotide and/or amino acid sequence characters shall conform to the code set forth in the tables in paragraphs (b)(1) and (b)(2) of this section. No code other than that specified in this section shall be used in nucleotide and amino acid sequences. A modified base or amino acid may be presented in a given sequence as the corresponding unmodified base or amino acid if the modified base or amino acid is one of those listed in

paragraphs (p)(1) or (p)(2) of this section and the modification is also set forth elsewhere in the Sequence Listing (for example, FEATURES §1.823(b)(2)(ix)). Otherwise, all bases or amino acids not appearing in paragraphs (b)(1) or (b)(2) of this section shall be listed in a given sequence as "N" or "Xaa," respectively, with further information, as appropriate, given elsewhere in the Sequence Listing.

(1) Base codes:

Symbol	Meaning
A	A; adenine
C	C; cytosine
G	G; guanine
T	T; thymine
U	U; uracil
M	A or C
R	A or G
W	A or T/U
S	C or G
Y	C or T/U
K	G or T/U
V	A or C or G; not T/U
H	A or C or T/U; not G
D	A or G or T/U; not C
B	C or G or T/U; not A
N	(A or C or G or T/U) or (unknown or other)

(2) Amino acid three-letter abbreviations:

Abbreviation	Amino acid name
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
Asx	Aspartic Acid or Asparagine
Cys	Cysteine
Glu	Glutamic Acid
Gln	Glutamine
Glx	Glutamine or Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine

Val	Valine
Xaa	Unknown or other

(c) A nucleotide sequence shall be listed using the one-letter code for the nucleotide bases, as in paragraph (b)(1) of this section.

(d) The amino acids corresponding to the codons in the coding parts of a nucleotide sequence shall be typed immediately below the corresponding codons. Where a codon spans an intron, the amino acid symbol shall be typed below the portion of the codon containing two nucleotides.

(e) The amino acids in a protein or peptide sequence shall be listed using the three-letter abbreviation with the first letter as an upper case character, as in paragraph (b)(2) of this section.

(f) The bases in a nucleotide sequence (including introns) shall be listed in groups of 10 bases except in the coding parts of a sequence. Leftover bases, fewer than 10 in number, at the end of noncoding parts of a sequence shall be grouped together and separated from adjacent groups of 10 or 3 bases by a space.

(g) The bases in the coding parts of a nucleotide sequence shall be listed as triplets (codons).

(h) A protein or peptide sequence shall be listed with a maximum of 16 amino acids per line, with a space provided between each amino acid.

(i) A nucleotide sequence shall be listed with a maximum of 16 codons or 60 bases per line, with a space provided between each codon or group of 10 bases.

(j) A nucleotide sequence shall be presented, only by a single strand, in the 5' to 3' direction, from left to right.

(k) An amino acid sequence shall be presented in the amino to carboxy direction, from left to right, and the amino and carboxy groups shall not be presented in the sequence.

(l) The enumeration of nucleotide bases shall start at the first base of the sequence with number 1. The enumeration shall be continuous through the whole sequence in the direction 5' to 3'. The enumeration shall be marked in the right margin, next to the line containing the one-letter codes for the bases, and giving the number of the last base of that line.

(m) The enumeration of amino acids may start at the first amino acid of the first mature protein, with number 1. The amino acids preceding the mature protein, e.g., pre-sequences, pro-sequences, pre-pro-sequences and signal sequences, when presented, shall have negative numbers, counting backwards starting with the amino acid next to number 1. Otherwise, the enumeration of amino acids shall start at the first amino acid at the amino terminal as number 1. It shall be marked below the sequence every 5 amino acids.

(n) For those nucleotide sequences that are circular in configuration, the enumeration method set forth in paragraph (l) of this section remains applicable with the exception that the designation of the first base of the nucleotide sequence may be made at the option of the applicant. The enumeration method for amino acid sequences that is set forth in paragraph (m) of this section remains applicable for amino acid sequences that are circular in configuration.

(o) A sequence with a gap or gaps shall be presented as a plurality of separate sequences, with separate sequence identifiers, with the number of separate sequences being equal in number to the number of continuous strings of sequence data. A sequence that is made up of one or more noncontiguous segments of a larger sequence or segments from different sequences shall be presented as a separate sequence.

(p) The code for representing modified nucleotide bases and modified and unusual amino acids shall conform to the code set forth in the tables in paragraphs (p)(1) and (p)(2) of this section. The modified base controlled vocabulary in paragraph (p)(1) of this section and the modified and unusual amino acids in paragraph (p)(2) of this section shall not be used in the nucleotide and/or amino acid sequences; but may be used in the description and/or the "Sequence Listing" corresponding to, but not including, the nucleotide and/or amino acid sequence.

(1) Modified base controlled vocabulary:

Abbreviation	Modified base description
ac4c	4-acetylcytidine
chm5u	5-(carboxyhydroxylmethyl)uridine
cm	2'-O-methylcytidine
cmnm5s2u	5-carboxymethylaminomethyl-2-thiouridine
cmnm5u	5-carboxymethylaminomethyluridine
d	dihydrouridine
fm	2'-O-methylpseudouridine
gal q	beta,D-galactosylqueosine
gm	2'-O-methylguanosine
i	inosine
i6a	N6-isopentenyladenosine
m1a	1-methyladenosine
m1f	1-methylpseudouridine
m1g	1-methylguanosine
m1i	1-methylinosine
m22g	2,2-dimethylguanosine
m2a	2-methyladenosine
m2g	2-methylguanosine
m3c	3-methylcytidine
m5c	5-methylcytidine
m6a	N6-methyladenosine
m7g	7-methylguanosine
mam5u	5-methylaminomethyluridine
mam5s2u	5-methoxyaminomethyl-2-thiouridine
man q	beta,D-mannosylqueosine
mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine
mcm5u	5-methoxycarbonylmethyluridine
mo5u	5-methoxyuridine
ms2i6a	2-methylthio-N6-isopentenyladenosine
ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl) carbamoyl)threonine
mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methyl-carbamoyl)threonine
mv	uridine-5-oxyacetic acid methylester
o5u	uridine-5-oxyacetic acid (v)

osyw	wybutoxosine
p	pseudouridine
q	queosine
s2c	2-thiocytidine
s2t	5-methyl-2-thiouridine
s2u	2-thiouridine
s4u	4-thiouridine
t	5-methyluridine
t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)carbamoyl) threonine
tm	2'-O-methyl-5-methyluridine
um	2'-O-methyluridine
yw	wybutosine
x	3-(3-amino-3-carboxypropyl)uridine, (acp3)u

(2) Modified and unusual amino acids:

Abbreviation	Modified and unusual amino acid
Aad	2-Aminoadipic acid
bAad	3-aminoadipic acid
bAla	beta-Alanine, beta-Aminopropionic acid
Abu	2-Aminobutyric acid
4Abu	4-Aminobutyric acid, piperidinic acid
Acp	6-Aminocaproic acid
Ahe	2-Aminoheptanoic acid
Aib	2-Aminoisobutyric acid
bAib	3-Aminoisobutyric acid
Apm	2-Aminopimelic acid
Dbu	2,4-Diaminobutyric acid
Des	Desmosine
Dpm	2,2'-Diaminopimelic acid
Dpr	2,3-Diaminopropionic acid
EtGly	N-Ethylglycine
EtAsn	N-Ethylasparagine
Hyl	Hydroxylysine
aHyl	allo-Hydroxylysine
3Hyp	3-Hydroxyproline
4Hyp	4-Hydroxyproline
Ide	Isodesmosine
aIle	allo-Isoleucine
MeGly	N-Methylglycine, sarcosine
MeIle	N-Methylisoleucine
MeLys	6-N-Methyllysine
MeVal	N-Methylvaline
Nva	Norvaline
Nle	Norleucine
Orn	Ornithine

§ 1.823 Requirements for nucleotide and/or amino acid sequences as part of the application papers.

(a) The "Sequence Listing," required by § 1.821(c), setting forth the nucleotide and/or amino acid sequences, and associated information in accordance with paragraph (b) of this section, must begin on a new page and be titled "Sequence Listing" and appear immediately prior to the claims. Each page of the "Sequence Listing" shall contain no more than 66 lines and each line shall contain no more than 72 characters. A fixed-width font shall be used exclusively throughout the "Sequence Listing."

(b) The "Sequence Listing" shall, except as otherwise indicated, include, in addition to and immediately preceding the actual nucleotide and/or amino acid sequence, the following items of information. The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below, except that parenthetical explanatory information following the headings (identifiers) is to be omitted. Each item of information shall begin on a new line, enumerated with the number/numeral/letter in parentheses as shown below, with the heading (identifier) in upper case characters, followed by a colon, and then followed by the information provided. Except as allowed below, no item of information shall occupy more than one line. Those items of information that are applicable for all sequences shall only be set forth once in the "Sequence Listing." The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "R" is recommended, but not required. The submission of those items of information designated with an "O" is optional. Those items designated with "rep" may have multiple responses and, as such, the item may be repeated in the "Sequence Listing."

(1) GENERAL INFORMATION (Application, diskette/tape and publication information):

(i) APPLICANT (maximum of first ten named applicants; specify one name per line: SURNAME comma OTHER NAMES and/or INITIALS - M/rep):

(ii) TITLE OF INVENTION (title of the invention, as elsewhere in application, four lines maximum - M):

(iii) NUMBER OF SEQUENCES (number of sequences in the "Sequence Listing" - M):

(iv) CORRESPONDENCE ADDRESS (M):

(A) ADDRESSEE (name of applicant, firm, company or institution, as may be appropriate):
(B) STREET (correspondence street address, as elsewhere in application, four lines maximum):
(C) CITY (correspondence city address, as elsewhere in application):
(D) STATE (correspondence state, as elsewhere in application):
(E) COUNTRY (correspondence country, as elsewhere in application):
(F) ZIP (correspondence zip or postal code, as elsewhere in

application):

(v) COMPUTER READABLE FORM (M):

(A) MEDIUM TYPE (type of diskette/tape submitted):
(B) COMPUTER (type of computer used with diskette/tape submitted):
(C) OPERATING SYSTEM (type of operating system used):
(D) SOFTWARE (type of software used to create computer readable form):

(vi) CURRENT APPLICATION DATA (M, if available):

(A) APPLICATION NUMBER (U.S application number, including a series code, a slash and a serial number, or U.S. PCT application number, including the letters PCT, a slash, a two letter code indicating the U.S. as the Receiving Office, a two digit indication of the year, a slash and a five digit number, if available):
(B) FILING DATE (U.S. or PCT application filing date, if available; specify as dd-MMM-yyyy):
(C) CLASSIFICATION (IPC/US classification or F-term designation, where F-terms have been developed, if assigned, specify each designation, left justified, within an eighteen position alpha numeric field - rep, to a maximum of ten classification designations):

(vii) PRIOR APPLICATION DATA (prior domestic, foreign priority or international application data, if applicable - M/rep):

(A) APPLICATION NUMBER (application number; specify as two letter country code and an eight digit application number, or if a PCT application, specify as the letters PCT, a slash, a two letter code indicating the Receiving Office, a two digit indication of the year, a slash and a five digit number):
(B) FILING DATE (document filing date, specify as dd-MMM-yyyy):

(viii) ATTORNEY/AGENT INFORMATION (O):

(A) NAME (attorney/agent name; SURNAME comma OTHER NAMES and/or INITIALS):
(B) REGISTRATION NUMBER (attorney/agent registration number):
(C) REFERENCE/DOCKET NUMBER (attorney/agent reference or docket number):

(ix) TELECOMMUNICATION INFORMATION (O):

(A) TELEPHONE (telephone number of applicant or attorney/agent):
(B) TELEFAX (telefax number of applicant or

attorney/agent):
(C) TELEX (telex number of applicant or attorney/agent):

(2) INFORMATION FOR SEQ ID NO: X (rep):

(i) SEQUENCE CHARACTERISTICS (M):

(A) LENGTH (sequence length, expressed as number of base pairs or amino acid residues):
(B) TYPE (sequence type, i.e., whether nucleic acid or amino acid):
(C) STRANDEDNESS (if nucleic acid, number of strands of source organism molecule, i.e., whether single stranded, double stranded, both or unknown to applicant):
(D) TOPOLOGY (whether source organism molecule is circular, linear, both or unknown to applicant):

(ii) MOLECULE TYPE (type of molecule sequenced in SEQ ID NO:X (at least one of the following should be included with subheadings, if any, in Sequence Listing - R)):
- Genomic RNA;
- Genomic DNA;
- mRNA
- tRNA;
- rRNA;
- snRNA;
- scRNA;
- preRNA;
- cDNA to genomic RNA;
- cDNA to mRNA;
- cDNA to tRNA;
- cDNA to rRNA;
- cDNA to snRNA;
- cDNA to scRNA;
- Other nucleic acid;
(A) DESCRIPTION (four lines maximum):
- protein and
- peptide.

(iii) HYPOTHETICAL (yes/no - R):

(iv) ANTI-SENSE (yes/no - R):

(v) FRAGMENT TYPE (for proteins and peptides only, at least one of the following should be included in the Sequence Listing - R):
- N-terminal fragment;
- C-terminal fragment and
- internal fragment.

(vi) ORIGINAL SOURCE (original source of molecule sequenced in SEQ ID NO:X - R):

(A) ORGANISM (scientific name of source organism):
(B) STRAIN:
(C) INDIVIDUAL ISOLATE (name/number of individual/isolate):
(D) DEVELOPMENTAL STAGE (give developmental stage of source organism and indicate whether derived from germ-line or rearranged developmental pattern):
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE (immediate experimental source of the sequence in SEQ ID NO:X - R):

(A) LIBRARY (library -type,name):
(B) CLONE (clone(s)):

(viii) POSITION IN GENOME (position of sequence in SEQ ID NO:X in genome - R):

(A) CHROMOSOME/SEGMENT (chromosome/segment - name/number):
(B) MAP POSITION:
(C) UNITS (units for map position, i.e., whether units are genome percent, nucleotide number or other/specify):

(ix) FEATURE (description of points of biological significance in the sequence in SEQ ID NO:X -R/rep):

(A) NAME/KEY (provide appropriate identifier for feature - four lines maximum):
(B) LOCATION (specify location according to syntax of DDBJ/EMBL/GenBank Feature Tables Definition, including whether feature is on complement of presented sequence; where appropriate state number of first and last bases/amino acids in feature - four lines maximum):
(C) IDENTIFICATION METHOD (method by which the feature was identified, i.e., by experiment, by similarity with known sequence or to an established consensus sequence, or by similarity to some other pattern - four lines maximum):
(D) OTHER INFORMATION (include information on phenotype conferred, biological activity of sequence or its product, macromolecules which bind to sequence or its product, or other relevant information - four lines maximum):

(x) PUBLICATION INFORMATION (Repeat section for each relevant publication - O/rep):

(A) AUTHORS (maximum of first ten named authors of publication; specify one name per line: SURNAME comma

OTHER NAMES and/or INITIALS - rep):
(B) TITLE (title of publication):
(C) JOURNAL (journal name in which data published):
(D) VOLUME (journal volume in which data published):
(E) ISSUE (journal issue number in which data published):
(F) PAGES (journal page numbers in which data published):
(G) DATE (journal date in which data published; specify as dd-MMM-yyyy, MMM-yyyy or Season-yyyy):
(H) DOCUMENT NUMBER (document number, for patent type citations only; specify as two letter country code, eight digit document number (right justified), one letter and, as appropriate, one number or a space as a document type code; or if a PCT application, specify as the letters PCT, a slash, a two letter code indicating the Receiving Office, a two digit indication of the year, a slash and a five digit number; or if a PCT publication, specify as the two letters WO, a two digit indication of the year, a slash and a five digit publication number):
(I) FILING DATE (document filing date, for patent-type citations only; specify as dd-MMM-yyyy):
(J) PUBLICATION DATE (document publication date; for patent-type citations only, specify as dd-MMM-yyyy):
(K) RELEVANT RESIDUES IN SEQ ID NO:X (rep): FROM (position) TO (position)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X:

§ 1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by § 1.821(e) shall contain a printable copy of the "Sequence Listing," as defined in §§ 1.821(c), 1.822 and 1.823, recorded as a single file on either a diskette or a magnetic tape. The computer readable form shall be encoded and formatted such that a printed copy of the "Sequence Listing" may be recreated using the print commands of the computer/operating-system configurations specified in paragraph (f) of this section.

(b) The file in paragraph (a) of this section shall be encoded in a subset of the American Standard Code for Information Interchange (ASCII). This subset shall consist of all the printable ASCII characters including the ASCII space character plus line-termination, pagination and end-of-file characters associated with the computer/operating-system configurations specified in paragraph (f) of this section. No other characters shall be allowed.

(c) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall be readable by one of the computer/operating-system configurations specified in paragraph (f) of this section, and shall conform to the specifications in paragraphs (a) and (b) of this section.

(d) The entire printable copy of the "Sequence Listing" shall be contained within one file on a single diskette or magnetic tape unless it is shown to the satisfaction of the Commissioner that it is not practical or possible to submit the entire printable copy of the "Sequence Listing" within

one file on a single diskette or magnetic tape.

(e) The submitted diskette or tape shall be write-protected such as by covering or uncovering diskette holes, removing diskette write tabs or removing tape write rings.

(f) As set forth in paragraph (c), above, any means may be used to create the computer readable form, as long as the following conditions are satisfied. A submitted diskette shall be readable on one of the computer/operating-system configurations described in paragraphs (1) through (3), below. A submitted tape shall satisfy the format specifications described in paragraph (4), below.

(1) Computer: IBM PC/XT/AT, IBM PS/2 or compatibles;
Operating system: PC-DOS or MS-DOS (Versions 2.1 or above);
Line Terminator: ASCII Carriage Return plus ASCII Line Feed;
Pagination: ASCII Form Feed or Series of Line Terminators;
End-of-File: ASCII SUB (Ctrl-Z);
Media: Diskette - 5.25 inch, 360 Kb storage;
Diskette - 5.25 inch, 1.2 Mb storage;
Diskette - 3.50 inch, 720 Kb storage;
Diskette - 3.50 inch, 1.44 Mb storage;
Print Command: PRINT filename.extension;

(2) Computer: IBM PC/XT/AT, IBM PS/2 or compatibles;
Operating system: Xenix;
Line Terminator: ASCII Carriage Return;
Pagination: ASCII Form Feed or Series of Line Terminators;
End-of-File: None;
Media: Diskette - 5.25 inch, 360 Kb storage;
Diskette - 5.25 inch, 1.2 Mb storage;
Diskette - 3.50 inch, 720 Kb storage;
Diskette - 3.50 inch, 1.44 Mb storage;
Print Command: lpr filename;

(3) Computer: Apple Macintosh;
Operating System: Macintosh;
Macintosh File Type: text with line termination
Line Terminator: Pre-defined by text type file;
Pagination: Pre-defined by text type file;
End-of-file: Pre-defined by text type file;
Media: Diskette - 3.50 inch, 400 Kb storage;
Diskette - 3.50 inch, 800 Kb storage;
Diskette - 3.50 inch, 1.4 Mb storage;
Print Command: Use PRINT command from any Macintosh Application that processes text files, such as MacWrite or TeachText;

(4) Magnetic tape: 0.5 inch, up to 2400 feet;
Density: 1600 or 6250 bits per inch, 9 track;
Format: raw, unblocked;
Line Terminator: ASCII Carriage Return plus optional ASCII Line Feed;
Pagination: ASCII Form Feed or Series of Line Terminators;
Print Command (Unix shell version given here as sample response -

mt/dev/rmt0; lpr</dev/rmt0):

(g) Computer readable forms that are submitted to the Office will not be returned to the applicant.

(h) All computer readable forms shall have a label permanently affixed thereto on which has been hand printed or typed, a description of the format of the computer readable form as well as the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form and the name and type of computer and operating system which generated the files on the computer readable form. If all of this information can not be printed on a label affixed to the computer readable form, by reason of size or otherwise, the label shall include the name of the applicant and the title of the invention and a reference number, and the additional information may be provided on a container for the computer readable form with the name of the applicant, the title of the invention, the reference number and the additional information affixed to the container. If the computer readable form is submitted after the date of filing under 35 U.S.C. 111, after the date of entry in the national stage under 35 U.S.C. 371 or after the time of filing, in the United States Receiving Office, an international application under the PCT, the labels mentioned herein must also include the date of the application and the application number, including series code and serial number.

§ 1.825 Amendments to or replacement of sequence listing and computer readable copy thereof.

(a) Any amendment to the paper copy of the "Sequence Listing" (§ 1.821(c)) must be made by the submission of substitute sheets. Amendments must be accompanied by a statement that indicates support for the amendment in the application, as filed, and a statement that the substitute sheets include no new matter. Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(b) Any amendment to the paper copy of the "Sequence Listing," in accordance with paragraph (a) of this section, must be accompanied by a substitute copy of the computer readable form (§ 1.821(e)) including all previously submitted data with the amendment incorporated therein, accompanied by a statement that the copy in computer readable form is the same as the substitute copy of the "Sequence Listing." Such a statement must be a verified statement if made by a person not registered to practice before the Office.

(c) Any appropriate amendments to the "Sequence Listing" in a patent, e.g., by reason of reissue or certificate of correction, must comply with the requirements of paragraphs (a) and (b) of this section.

(d) If, upon receipt, the computer readable form is found to be damaged or unreadable, applicant must provide, within such time as set by the Commissioner, a substitute copy of the data in computer readable form accompanied by a statement that the substitute data is identical to that originally filed. Such a statement must be a verified statement if made by a person not registered to practice before the Office.